

Tilburg University

The maximum number of omitted variables, Problem 00.2.2

Magnus, J.R.; Danilov, D.L.

Published in:
Econometric Theory

Publication date:
2000

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):
Magnus, J. R., & Danilov, D. L. (2000). The maximum number of omitted variables, Problem 00.2.2. *Econometric Theory*, 16, 287-288.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

00.2.2. *The Maximum Number of Omitted Variables*—Solution, proposed by Dmitri L. Danilov and Jan R. Magnus. If $r = 0$, the result is simple and well known. We assume that $r \geq 1$. Let $(S: T)$ be an orthogonal $k_2 \times k_2$ matrix such that

$$X_2' X_1 X_1' X_2 S = S\Lambda, \quad X_1' X_2 T = 0,$$

where Λ denotes an $r \times r$ diagonal matrix with positive diagonal elements. Notice that the dimensions of S and T are $k_2 \times r$ and $k_2 \times (k_2 - r)$, respectively. Because $X_2' X_2$ has full rank k_2 , we obtain

$$r(X_2 T) = r(X_2' X_2 T) = r(T) = k_2 - r,$$

so that the $n \times (k_2 - r)$ matrix $W_2 \equiv X_2 T$ has full column rank. Hence, we may define the idempotent matrix $M_2 = I_n - W_2(W_2' W_2)^{-1} W_2'$.

Now, let $W_1 \equiv M_2 X_2 S$, an $n \times r$ matrix. Because

$$W_1 = M_2 X_2 S = X_2 S - W_2(W_2' W_2)^{-1} W_2' X_2 S,$$

we obtain $X_1' W_1 = X_1' X_2 S$ and, hence,

$$X_2' X_1 X_1' W_1 = X_2' X_1 X_1' X_2 S = S\Lambda,$$

so that $r = r(X_2' X_1 X_1' W_1) \leq r(W_1) \leq r$ and, hence, $r(W_1) = r$.

Next, let $W \equiv (W_1: W_2)$. We already know that $r(W_1) = r$ and $r(W_2) = k_2 - r$. Because $M_2 W_2 = 0$, it follows that $W_1' W_2 = 0$ and, hence, that $r(W) = r(W_1) + r(W_2) = r + k_2 - r = k_2$.

Finally, we observe that

$$M_2 X_2 = X_2 - X_2 T(W_2' W_2)^{-1} W_2' X_2 = X_2 P$$

for some matrix P and, hence,

$$W = (W_1: W_2) = (M_2 X_2 S: X_2 T) = (X_2 P S: X_2 T) = X_2 Q$$

for some $k_2 \times k_2$ matrix Q . Because $r(W) = k_2$, Q is non-singular.

It is now easy to see that W_2 is orthogonal to both W_1 and X_1 . Also, the space spanned by the k_2 columns of W is identical to the space spanned by the k_2 columns of X_2 , so that $X_2 \beta_2 = W \delta$ for some choice of δ (namely $\delta = Q^{-1} \beta_2$). Hence, the estimator $\hat{\beta}_1$ obtained from a regression of y on X_1 and X_2 will be identical to the estimator obtained from a regression of y on X_1 and W_1 , and W_1 only has r columns.

When drawing inferences about β_1 , we assume that $u \sim N(0, \sigma^2 I_n)$. The estimator of σ^2 will be biased upward if we delete W_2 from our regression, even though W_2 is orthogonal to both X_1 and W_1 , just as in the standard textbook case.